

524,566

(12) INTERNATIONAL PUBLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
26 February 2004 (26.02.2004)

PCT

(10) International Publication Number
WO 2004/017201 A2

- (51) International Patent Classification⁷: **G06F 9/46**
- (21) International Application Number:
PCT/EP2003/010081
- (22) International Filing Date: 13 August 2003 (13.08.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
02368088.7 13 August 2002 (13.08.2002) EP
- (71) Applicant (for all designated States except MC, US):
INTERNATIONAL BUSINESS MACHINES CORPORATION [US/US]; New Orchard Road, Armonk, NY 10504 (US).
- (71) Applicant (for MC only): **COMPAGNIE IBM FRANCE** [FR/FR]; Tour Descartes - La Defense 5, 2, avenue Gambetta, F-92400 Courbevoie (FR).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **DONATELLI, Alessandro** [IT/IT]; Via della Magliana Nuova, 150,

I-00146 Roma (IT). **D'ALO, Salvatore** [IT/IT]; Via Nino Taranto, 30, Fabbriato A4, Scala B, Interno 13, I-00125 Roma (IT). **LANFRANCHI, Giovanni** [IT/IT]; Via De Longe, 10, I-29100 Piacenza (IT). **MARINELLI, Claudio** [IT/IT]; Via Goito n°6, I-04011 Aprilia (IT).

(74) Agent: **ZERBI, Guido**; Compagnie IBM France, Direction de la Propriété Intellectuelle, F-06610 La Gaude (FR).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

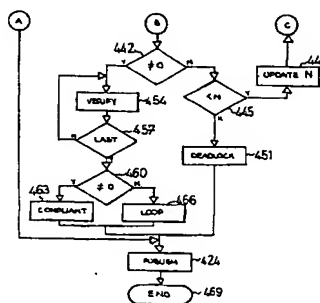
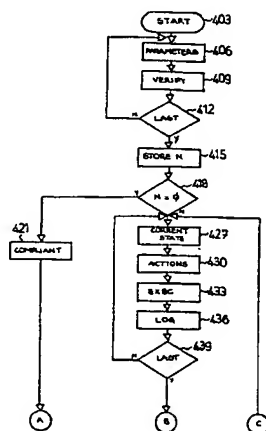
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv)) for US only

[Continued on next page]

(54) Title: AN ADAPTIVE MANAGEMENT METHOD AND SYSTEM WITH AUTOMATIC DEPENDENCY RESOLUTION



(57) Abstract: A resource management method (300) and a corresponding system are proposed. In the solution of the invention, an authority publishes (324) multiple rules, each one defining a desired target configuration for a category of subjects (without any information about their dependencies). Each subject retrieves (320-322, 326) the rules corresponding to its category in a random order. The rules are then applied (427-439) on the subject according to a trial-and-fail approach. Particularly, the application of any failed rule is continually repeated (427-448), until all the rules are successfully applied (463) or a deadlock condition is detected (451). In this way, any dependency is automatically resolved on the subject at run time. Moreover, as soon as all the rules have been successfully applied the compliance of the subject to the rules is verified (455-57) again; should the subject be not compliant to one or more rules any longer, an infinite loop condition is detected (466).

WO 2004/017201 A2



Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**AN ADAPTIVE MANAGEMENT METHOD AND SYSTEM WITH AUTOMATIC
DEPENDENCY RESOLUTION**

Technical field

The present invention relates to the data processing
5 field, and more specifically to a resource management method
and a corresponding system.

Background art

Management of different types of resources (such as
software components, applications or devices) is a critical
10 issue in a data processing system with a distributed
architecture. This problem is particular acute when the system
includes a high number of logic and/or physic entities
(referred to as subjects), each one directly controlling one
or more resources; the problem is further exacerbated if the
15 subjects have a high level of complexity or are dispersed
across a large number of installations.

The management environments known in that art are based on
an enforcement model (also known as manager/workers model). In
this model, the process is entirely controlled by an authority
20 residing at a central site of the system. The authority
defines a desired configuration of every subject. For this

purpose, the authority accesses a central repository storing the (alleged) current configuration of each subject, and determines the management actions required to bring the subject to the desired configuration starting from the current
5 configuration. The management actions are then enforced remotely by the authority on the subject (which is totally passive).

A drawback of the management environments known in the art is the lack of any kind of cooperation between the
10 authority and the subjects. This lack of cooperation may bring about inconsistencies when the subjects are upgraded out of the control of the authority. Moreover, in the solutions currently employed the management of subjects that are temporarily unavailable or off-line is quit difficult to
15 implement. The known management environments require the authority to maintain information about the location of all the subjects; at the same time, the authority must handle the communication with every subject directly.

An additional problem arises when a specific state of one
20 or more resources is a pre-requisite to the execution of some management actions. A typical example is that of a patch for a software product; in this case, the installation of the patch requires that the software product be already available on the subject. More complex dependencies are experienced in large
25 systems with several correlated tiers.

As a consequence, the authority must explicitly define a workflow that drives the enforcement of the management actions on each subject in the correct sequence.

The authority must also handle any conditioning among the
30 management actions directly. In any case, some conflicts cannot be identified a priori. Conversely, those conflicts arise if particular (and often unforeseeable) conditions occur; therefore, they can be detected only at run time when the management actions are enforced (assuming that the
35 subjects report the result of the enforcement to the authority).

This problem is particular acute in high dynamic environments, wherein complex dependencies are defined among the different resources.

Moreover, the above-mentioned drawbacks strongly increase the difficulty of correctly defining any solution to be deployed in the system.

Summary of the invention

It is an object of the present invention to provide a management method that implements a self-adaptive model.

10 Particularly, it is an object of the present invention to resolve any dependency among the management actions automatically.

It is another object of the present invention to ensure a correct sequencing of the management actions on each subject, without the need of defining a corresponding workflow explicitly.

It is yet another object of the present invention to handle any conditioning in the management actions automatically.

20 Moreover, it is an object of the present invention to simplify the definition of any solution to be deployed in the system.

The accomplishment of these and other related objects is achieved, in a data processing system including a plurality of subject entities and at least one authority entity, by an autonomic management method for self-configuring the subject entities, each subject entity belonging to at least one of a plurality of categories, wherein the method includes the steps of: the at least one authority entity publishing a plurality of rules each one defining a target state for a category, the target state of at least one rule being a prerequisite for the target state of at least one further rule, each subject entity retrieving the rules for the corresponding at least one category, applying each retrieved rule to configure the

subject entity according to the target state, the application of the retrieved rule failing when at least one corresponding prerequisite is not available on the subject entity, and repeating the application of each failed rule to configure the
5 subject entity according to the target state specified in the failed rule when all the corresponding prerequisites are available.

The present invention also provides a computer program for performing the method and a corresponding computer program
10 running on each subject entity; program products storing those programs are also encompassed.

Moreover, the present invention provides a system implementing the method and a computer of the subject entity for use in the system.

15 The novel features believed to be characteristic of this invention are set forth in the appended claims. The invention itself, however, as well as these and other related objects and advantages thereof, will be best understood by reference to the following detailed description to be read in
20 conjunction with the accompanying drawings.

Brief description of the drawings

Figure 1a	is a schematic block diagram of a data processing system in which the management method of the invention is applicable;
Figure 1b	shows the functional blocks of a generic computer of the system;
Figure 2	depicts the main software components used for implementing the management method;
Figure 3	is a diagram describing the flow of activities relating to the self-configuration of a subject in the system;
Figures 4a-4b	show a flow chart specifying the actions executed on the subject;
Figure 5	depicts a class diagram implementing the management method; and
Figure 6	is a sequence diagram exemplifying a specific management operation.

Detailed description of the preferred embodiment

With reference in particular to Figure 1a, a data processing system 100 with a distributed architecture (typically INTERNET-based) is shown. The system 100 implements
5 an environment for managing several kinds of resources, such as products in a software distribution application, in a monitoring application, and the like.

A computer 105 operates as an authority, which is responsible to define a desired configuration of the system
10 100. Multiple computers 110 operate as heterogeneous subjects, which directly control one or more instances of different resources under management. Each subject 110 is defined by one or more characteristics of logic and/or physic entities (such as a desktop, a Personal Digital Assistant (PDA), a SIM card
15 of a mobile telephone, every user of a workstation, and the like).

The authority 105 and the subjects 110 communicate in an associative way (at the information level) through a shared repository 115. The repository 115 stores a set of rules (or
20 laws). Each rule specifies a desired configuration for a category of subjects, defined according to their logic and/or physic characteristics; particularly, the desired configuration is characterized by a target state of one or more resources (which instances are controlled by the subjects
25 belonging to the category associated with the rule). Moreover, the rule can include formal parameters (or arguments), which are used to pass a value or a reference dynamically when the rule is applied on the corresponding subjects; for example, the formal parameters define correlations with different
30 categories.

The rules are established regardless of any dependencies among the target states of the respective resources (which dependencies are resolved at run time, as described in detail

in the following). In other words, each rule does not include any information about the target states of one or more resources that are pre-requisite for the target state of the respective resource. For example, when the installation of a software product is required for installing a corresponding patch (on the subjects belonging to a specific category), two rules are issued simply specifying that the software product and the patch, respectively, must be installed for that category.

10 A computer 120 operates as an administrator, which is responsible to define the categories of the subjects 110 dynamically at a logic level. The information is available in a shared repository 125; the relevant information stored in the logic data repository 125 is transmitted to every subject
15 110.

 The subjects 110 communicate through an additional shared repository 130 (defining a virtual word implemented in the system 100); for each subject 110, the repository 130 stores membership information indicating the category of the subject
20 and its compliance to the corresponding rules. The membership data repository 130 is further accessed by the administrator 120.

 The above-described repositories implement a peer-to-peer communication model, which totally decouples the different
25 entities of the system 100 (authority 105, subjects 110 and administrator 120). Particularly, a destination decoupling results from the fact that the entities do not need to refer to each other explicitly (thereby providing a fully anonymous communication scheme); in other words, a mutual knowledge of
30 the location of the different entities is not necessary. Moreover, time decoupling results from the fact that the entities do not need to be available at the same time.

 As shown in Figure 1b, a generic computer of the system (authority, subject, or administrator) is formed by several
35 units that are connected in parallel to a communication bus 145. In detail, one or more microprocessors (μ P) 150 control

operation of the computer, a DRAM 155 is directly used as a working memory by the microprocessor(s) 150, and a ROM 160 stores basic code for a bootstrap of the computer. Several peripheral units are further connected to the bus 145 (by means of respective interfaces). Particularly, a mass memory consists of a magnetic hard-disk 165 and a driver 170 for reading CD-ROMs 175. Moreover, the computer includes input devices 180 (for example, a keyboard and a mouse), and output devices 185 (for example, a monitor and a printer). A network Interface Card (NIC) 190 is used to connect the computer in the system.

However, the concepts of the present invention are also applicable when the management environment is implemented in an equivalent data processing system (for example, based on a LAN), or when different resources and entities are envisaged. Similar considerations apply if the computers have another architecture or include equivalent units. Alternatively, two or more authorities and/or administrators are provided, the authority and the administrator consist of a single entity, or different repositories are used for publishing the information about the categories of the subjects and the information about the compliance of the subjects to the corresponding rules, respectively.

Considering now Figure 2, the main software components used in the management environment described above are illustrated. The information (programs and data) is typically stored on the hard-disks of the computers and loaded (at least partially) into their working memories when the programs are running, together with an operating system and other application programs (not shown in the figure). The programs are initially installed onto the hard-disks from CD-ROMs.

Particularly, the authority 105 includes a ruling module 203 that is used to define new rules (for example, implementing a solution to be deployed in the system). Each new rule is validated for its semantic correctness; if the new

rule is not in contrast to the other rules, the module 203 publishes the new rule into the corresponding repository 115.

The rule repository 115 is implemented as a Web application. In detail, a relational database 209 stores all the rules currently in force. An engine 212 controls the execution of different management operations on the database 209 (for example, insertion of new rules, deletion of pre-existing rules or extraction of selected rules). In addition, each subject 110 can register with the engine 212; for this purpose, the subject 110 submits an identifier (typically consisting of a Globally Unique Identifier, or GUID) and an indication of its category; the information about the registered subjects is stored in a catalog 215 controlled by the engine 212.

Each subject 110 includes a membership controller 218, which assigns the subject 110 to the respective category (according to its logic and/or physic characteristics). For this purpose, the membership controller 218 cooperates with one or more plug-in modules.

Particularly, a first module 221 assigns the subject 110 to a category defined according to its physic properties. For example, the physic category is specified by hardware characteristics (such as a hard-disk size, a CPU model, or a working memory capacity) and/or software characteristics (such as installed applications, files or folders). For this purpose, the module 221 leverages a hardware inventory scanner, a software inventory scanner and an application scanner (not shown in the figure).

A different module 224 statically assigns the subject 110 to a category defined according to its logic properties. For example, the static logic category specifies different groups of users (such as secretaries, managers, developers, system engineers) or different characteristics of physic entities (such as desktops, laptops, PDAs, mobile telephones); the static logic category of the subject 110 is derived from an

identification code that is input during a log-in procedure or is hardwired.

Another module 227 dynamically assigns the subject 110 to a category defined according to further logic properties. For example, the dynamic logic category specifies a current function of the subject 110 in the system (such as server or agent in a monitoring application, depot in a software distribution application, and the like).

The dynamic logic categories of the different subjects 110 are set by a configuration module 230 running on the administrator 120. For this purpose, the configuration module 230 accesses an inventory 233 (for example, storing information about the organization of a company wherein the system is implemented). The information specifying the dynamic logic categories assigned to the subjects 110 is then published into the corresponding repository 125.

The logic data repository 125 (implemented as a Web application) includes a relational database 236 storing the category information; an engine 239 controls the database 236. Each subject 110 (identified by the corresponding GUID) can register with the engine 239; the information about the registered subjects 110 is stored in a catalog 242 controlled by the engine 239. The module 227 directly interfaces with the engine 239, in order to retrieve the information defining the dynamic logic category of the subject 110.

The membership controller 218 drives a membership provider 245. The membership provider 245 publishes information specifying the category of the subject (physic category, static logic category and/or dynamic logic category) into the corresponding repository 130. The membership data repository 130 (implemented as a Web application) includes a relational database 248 storing the membership information; an engine 251 controls the database 248. In addition, the engine 251 also interfaces with the configuration module 230 running on the administrator 120.

The membership controller 218 further supplies the information identifying the category of the subject 110 to a compliance engine 254. The compliance engine 254 retrieves the rules associated with the category of the subject 110 from the 5 corresponding repository 115 (through the engine 212). The retrieved rules are stored into a log 257 and then applied on the subject 110; for each rule, the log 257 also stores a flag indicating whether the subject is compliant or not to the rule. The compliance engine 254 interfaces with a class loader 10 260, which is used to load the classes 263 of the required resources on the subject 110 (from one or more repositories, not shown in the figure). Each resource class 263 controls a state catalog, which specifies the current state of each resource under management. The resource class 263 further 15 accesses a transition table; for each resource controlled by the subject 110 and for each pair current state/target state, the transition table indicates one or more management actions required for bringing the resource from the current state to the target state.

20 The compliance engine 254 also leverages the membership provider 245, in order to identify the subjects belonging to the category specified in any formal parameter (included in the rules to be applied on the subject 110). In addition, the engine 254 supplies information, indicative of the compliance 25 of the subject 110 to the corresponding rules, to the membership provider 245. The information is published into the membership data repository 130 (under the control of the engine 251).

Similar considerations apply if a whole application 30 (consisting of the programs on the different computers) and the corresponding data are structured in a different manner, if other modules or functions are provided, or if the programs are distributed on any other computer readable medium (such as one or more floppy-disks). However, the concepts of the present 35 invention are also applicable when the new rules are not validated by the authority, or when the shared repositories

are implemented with different technologies. Alternatively, the categories are based on other characteristics of the subjects, the physic and/or logic characteristics of the subjects are detected in a different manner, or the categories
5 are defined in another way (for example, with a different number and/or type of components, down to a single one).

In the system described above, each subject self-adapts to the rules published by the authority according to a trial-and-fail approach. As shown in Figure 3, a corresponding
10 process 300 begins at the black start circle 302 (in the swim-lane of the subject). Descending into block 304, the subject derives its static logic category from the identification code that is input during the log-in procedure or is hardwired.

15 Hardware and software scans are then executed at block 306 (assuming that the rules for installing the corresponding modules have already been applied), in order to detect the physic category of the subject; the same activities are also performed in response to the notification of a software and/or
20 hardware upgrade carried out on the subject.

The process continues to block 308, wherein the subject queries the logic data repository in order to identify its dynamic logic category. In response thereto, the engine of the logic data repository at block 310 retrieves the corresponding
25 information from its database; the category information is then returned to the subject. The block 310 is also reached from block 312 (in the swim-lane of the logic data repository), whenever updated information is published in the logic data repository for the subject (under the control of
30 the administrator); in this way, the subject is promptly notified of any change in its dynamic logic category (assuming that the subject has previously registered with the logic data repository). Returning to the swim-lane of the subject, the category information is received from the corresponding
35 repository at block 314.

Considering now block 316, information identifying the category of the subject (static logic category, physic category and dynamic logic category) is supplied to the membership data repository. The same activity is also performed whenever a predetermined time-out expires (for example, every hour). If the category has changed since a last verification, the information about the new category of the subject is published in the database of the membership data repository at block 318.

10 Referring back to the swim-lane of the subject, the rules corresponding to the category of the subject are requested to the rule repository at block 320. In response thereto, the engine of the rule repository at block 322 retrieves the relevant rules from its database; the rules are then returned
15 to the subject, so as to implement a pull mode of operation. The block 322 is also reached from block 324 (in the swim-lane of the rule repository), whenever new rules are published for the category of the subject (under the control of the authority); in this way, the subject is promptly notified of
20 any new rule applying to its category (assuming that the subject has previously registered with the rule repository), so as to implement a reactive mode of operation.

The process proceeds to block 326 in the swim-lane of the subject, wherein the rules corresponding to the category of
25 the subject are received from the respective repository. The received rules are then logged on the subject at block 328. Considering now block 330, all the rules are applied on the subject (as described in detail in the following); the same activity is also performed whenever a predetermined time-out
30 expires (for example, every day). The process then ends at the concentric white/black stop circles 332. In this way, the new rules are applied on the subject as soon as they are received; in addition, all the rules are periodically re-applied, so as to implement a healing mode of operation.

35 Moving now to Figures 4a-4b, the application of the rules on the subject implements a process that expands the block 330

described above. The process is executed whenever new rules for the subjects are retrieved from the respective repository (in response either to their publication or to the time-out expiration). Should the rules be received when the process is
5 already in progress, that process is aborted and a new process is started.

The activity flow of the process starts at block 403, and then passes to block 406. For each rule (starting from the first one), any formal parameter is resolved; for this purpose, the
10 formal parameter is replaced by the subject(s) belonging to the corresponding category (retrieved from the membership data repository). Referring now to block 409, the subject verifies whether the resource specified in the rule is already in the desired target state; if so, the compliance flag associated
15 with the rule is set. At the same time, in the event of any inconsistency of the resource, the corresponding entry of the state catalog is updated accordingly; this feature is used by the subject to self-repair its configuration (restoring the correct target state of any resource). For example, let us
20 assume that a component of an application installed on the subject (such as a file) has been removed by mistake or during installation of another application. In this case, the verification of the rule used to install the application detects the inconstancy and sets the state of the application
25 accordingly; as a consequence, when the rule is re-applied, the actions required to restore the missing file of the application will be performed automatically. Considering now block 412, the subject verifies whether the last rule has been processed. If not, the flow of activity returns to block 406
30 in order to execute the operations described above on a next rule.

As soon as all the rules have been processed, the method descends into block 415; the number of pending rules to which the subject is not compliant (denoted as N) is stored on the
35 subject. A test is then made at block 418 to verify whether the number of pending rules N is zero. If so, the process

passes to block 421, wherein a variable specifying the state of the subject is set to a value indicating its compliance to all the rules. The flow of activity then continues to block 424 (described in the following).

5 Returning to block 418, if the number of pending rules N is greater than zero the process enters an inner loop. For each pending rule (starting from the first one), the loop starts at block 427 wherein the current state of the corresponding resource is detected. Proceeding to block 430, a
10 list of the actions needed to bring the resource to the target state (specified in the pending rule) from the current state is extracted from the transition table. The process continues to block 433, wherein the actions are executed on the subject. Considering now block 436, the entry of the state catalogue
15 for the resource is updated accordingly. Moreover, if the application of the rule has been successful (bringing the resource to the desired target state) the corresponding compliance flag is set. The subject then verifies at block 439 whether the last pending rule has been processed. If not, the
20 process returns to block 427 in order to execute the operations described above on a next pending rule.

As soon as all the pending rules have been processed, the flow of activity descends into block 442; a test is made to determine whether the number of rules still pending is zero.
25 Is not, the subject verifies at block 445 whether the number of rules currently pending is lower than N (i.e., whether this number is decreased). If the result of the test is positive, the method continues to block 448 wherein the number of pending rules N is updated accordingly. The flow of activity
30 then returns to block 427, in order to repeat the inner loop 427-439.

Conversely, if the subject at block 445 determines that the number of pending rules N has not decreased the process descends into block 451. In this condition, a deadlock is
35 detected, since some pending rules are prevented from successfully bringing the corresponding resources to their

target states; in response thereto, the state variable is set accordingly to a value indicative of the deadlock condition. The process then continues to block 424.

For example, let us consider a rule specifying that a
5 corresponding resource must be installed on the subjects belonging to a specific category. The installation of the resource requires the fulfillment of some conditions (such as a minimum memory space available on the hard-disk). If those conditions are not met on a generic subject of the category,
10 the installation of the resource fails and the rule then hangs.

Referring back to block 442, if the number of pending rules N is zero the flow of activity descends into block 454. For each rule (starting from the first one), the subject
15 verifies again whether the resource specified in the rule is still in the desired target state; if not, the compliance flag associated with the rule is reset. Considering now block 412, if the last rule has been processed the flow of activity returns to block 406 in order to execute the operations
20 described above on a next rule. As soon as all the rules have been verified again, the process continues to the test block 460.

If the subject determines that it is still compliant to all the rules, the state variable is set accordingly.
25 Conversely, if the subject is not compliant to one or more rules any longer an infinite loop is detected, since the application of some rules impairs the target state reached by the resources of other rules; the state variable is set accordingly to a value indicative of the loop condition. In
30 both cases, the process then continues to block 424.

For example, let us consider a rule specifying that a resource must be installed on the subjects of a specific category; a further rule specifies that the same resource must be removed in the whole system. If the two rules are applied
35 in that order, the resource is installed on the subject by the first rule; the next application of the second rule removes

the resource, so that the subject will be no longer compliant to the first rule. Likewise, if the rules are applied in the reverse order, the resource is removed from the subject (if necessary) by the second rule; the next application of the first rule installs the resource, so that the subject will be no longer compliant to the second rule. In this condition, the application of the rules would repeat continually; in fact, the process never reaches a condition (terminating the loop) in which the subject is compliant to both the rules at the same time.

Considering now block 424, information about the compliance of the subject to the rules is transmitted to the membership data repository. Particularly, the state variable indicates whether the subject is compliant to all the rules (i.e., whether their application has been completed successfully or not); in the latest case, the state variable indicates whether the error is due to a deadlock or to an infinite loop. Moreover, for each rule the subject also provides information indicating whether the corresponding resource controlled by the subject is in the target state defined in the rule or not. The process then ends at the final block 469.

Similar considerations apply if an equivalent process is executed or if additional functions are provided. However, the concepts of the present invention are also applicable when the subjects operate in another way (for example, only in the pull mode or only in the reactive mode), or when the application of the rules on the subjects involves different operations. In alternative embodiments of the present invention, the rules do not support the definition of formal parameters, or the membership data repository is cleaned-up periodically; moreover, the administrator can enforce correction actions on non-compliant subjects, or the system can include one or more subjects operating as helpers for controlling the application of the rules on subjects that are not completely autonomous (for example, on mobile telephones).

Preferably, each rule is specified in the Extensible Markup Language (XML), which allows the creation of customized tags for any desired purpose. Each rule starts with a <LAW> tag and ends with a </LAW> tag. The rule includes a first part
5 (WHO clause), which is enclosed between the tags <WHO> and </WHO>, and a second part (WHAT clause), which is enclosed between the tags <WHAT> and </WHAT>.

The WHO clause defines the category associated with the rule by means of one or more key/value pairs. Each component
10 of the category is specified by a tag including the key GROUP, which defines the name of the category component; this tag is followed by the actual value of the category component enclosed between the tags <constant> and </constant>. The category components can be combined with standard logic
15 operators (such as AND, OR, NOT, and the like).

The WHAT clause describes the resource to be configured by the rule. In detail, the resource is defined between the tags <RESOURCE_DEFINITION> and </RESOURCE_DEFINITION>; for example, the rule specifies a type of the resource (enclosed
20 between the tags <RESOURCE_TYPE> and </RESOURCE_TYPE>), a repository from which corresponding software components can be downloaded (enclosed between the tags <RESOURCE_REPOSITORY> and </RESOURCE_REPOSITORY>), the name of its class (enclosed between the tags <RESOURCE_CLASS> and </RESOURCE_CLASS>), the
25 identification of a specific instance of the resource (enclosed between the tags <INSTANCE_IDENTIFIER> and </INSTANCE_IDENTIFIER>), and the like. The target state of the resource under management is defined between the tags <DESIRED_STATE> and </DESIRED_STATE>. Any argument of the rule
30 is enclosed between the tags <ARGUMENT> and </ARGUMENT>; the arguments can include one or more formal parameters, each one consisting of the name of an operational feature preceded by the symbol "\$", which name is followed by a dot and an expression identifying a category.

35 For example, the following rule ("RULE1"):

<LAW>

```
<WHO>
    <eq>
        <GROUP name="HD" />
        <constant>MyHD</constant>
5    </eq>
```

```
</WHO>
<WHAT>
    <RESOURCE_DEFINITION>
        <RESOURCE_TYPE>Program</RESOURCE_TYPE>
10    <INSTANCE_IDENTIFER>MyProgram</INSTANCE_IDENTIFIER>
    </RESOURCE_DEFINITION>
    <DESIRED_STATE>Installed</DESIRED_STATE>
</WHAT>
```

</LAW>

15 applies to all the subjects belonging to the category "HD=MyHD". The rule forces the subjects to have the instance "MyProgram" of the resource "Program" installed.

Likewise, the following further rule ("RULE2"):

```
<LAW>
20 <WHO>
    <eq>
        <GROUP name="HD" />
        <constant>MyHD</constant>
    </eq>
```

```
25 </WHO>
    <WHAT>
        <RESOURCE_DEFINITION>
            <RESOURCE_TYPE>Patch</RESOURCE_TYPE>
            <INSTANCE_IDENTIFER>MyPatch</INSTANCE_IDENTIFIER>
30    </RESOURCE_DEFINITION>
        <DESIRED_STATE>Installed</DESIRED_STATE>
    </WHAT>
```

</LAW>

forces the same subjects to have the instance "MyPatch" of the
35 resource "Patch" installed; the rule RULE2 requires the installation of the resource "MyProgram" as a pre-requisite; however, that dependency is not specified in the rule RULE2 in any way.

The above-described management environment is preferably implemented with a software application that is written in an object-oriented language (such as Java). As shown in Figure 5, a corresponding class diagram 500 includes a class "SUBJECT" 510 (representing a generic subject in the system). The class "SUBJECT" 510 instantiates objects of a class "RULE" 515, which represents a generic rule published by the authority. The class "RULE" 515 has a private attribute "RES_TYPE" (specifying the class name of the corresponding resource) and a private attribute "TG_STATE" (specifying the target state of the resource). Moreover, the class "RULE" 515 exposes a public method "APPLY" for implementing the rule on the subject; a private method "RESOLVE" is used to replace any formal parameter of the rule with the corresponding actual value.

The class "RULE" 515 has a collection consisting of any number of objects instantiating a class "PROPERTY" 520; each one of these objects includes additional information needed to apply the rule. A class "XML_RULE" 525 extends the class "RULE" 515; the class "XML_RULE" 525 defines an actual implementation of the rule.

The class "RULE" 515 and the class "SUBJECT" 510 instantiate objects implementing an interface "RESOURCE_TYPE" 530, which represents a generic resource under management. The interface "RESOURCE_TYPE" 530 declares an abstract attribute "TR_TABLE" for the corresponding transition table; moreover, the interface "RESOURCE_TYPE" 530 exposes a public method "VERIFY" (for validating the state of the resource) and a public method "SET_STATE" (for bringing the resource to the desired target state).

The interface "RESOURCE_TYPE" 530 is implemented by concrete classes 535 (generically denoted as "SPECIFIC_RESOURCE"), each one corresponding to a different type of resource (such as a file, a directory, a software package, a monitoring activity, and so on). Each class "SPECIFIC_RESOURCE" 535 defines the content of the transition table in the attribute "TR_TABLE".

The interface "RESOURCE_TYPE" 530 is further associated with any number of objects implementing an interface "MANAGEMENT_ACTION" 540 (which in turn is recursively associated with itself). The interface "MANAGEMENT_ACTION" 540
5 exposes a public method "EXEC" (for performing the required operations on the subject).

The interface "MANAGEMENT ACTION" 540 is implemented by concrete classes 545 (generically denoted as "SPECIFIC_ACTION"), each one corresponding to a different type
10 of management action (such as add, remove or change the attributes of a file/directory, install, remove or restore a software package, trigger or stop a monitoring activity, and so on). Each class "SPECIFIC_ACTION" 545 actually defines the method "EXEC".

15 The application of a generic rule involves the instantiation of an object of the type "XML_RULE" (passing the corresponding XML definition as a parameter), followed by the calling of the method "APPLY". At first, the method "APPLY" calls the method "RESOLVE" (in order to substitute the actual
20 values for any formal parameters of the rule). The method "APPLY" then instantiates an object of the type "SPECIFIC_RESOURCE" (whose name is stored in the attribute "RES_TYPE").

The verification of the compliance of the subject to the
25 rule (at the beginning and at the end of the application of all the rules) is carried out calling the method "VERIFY". The actual application of the rule starts with the calling of the method "SET_STATE" on the object of the type "SPECIFIC_RESOURCE" (passing the target state stored in the
30 attribute "TG_STATE" as a parameter). The method "SET_STATE" determines the current state of the resource and then queries the transition table (stored in the attribute "TR_TABLE"), in order to determine the management actions required to bring the resource to the target state from its current state. One
35 or more objects of the corresponding type "SPECIFIC_ACTION"

are instantiated, and the method "EXEC" is called on every one of them.

However, the concepts of the present invention are also applicable when the rules are defined in a different way (for example, using a tuple space), or when the software application implementing the management environment is written in a different language (even a non-object-oriented one). Similar considerations apply if the rules have a different structure, if other classes, interfaces, attributes and/or methods are envisaged, and the like.

An example of a scenario involved by the application of the two rules described in the foregoing ("RULE1" and "RULE2") is illustrated in the sequence diagram of Figure 6. A generic subject belonging to the category "MyHD" receives the rules "RULE1" and "RULE2" in a random order (for example, first the rule "RULE2" and then the rule "RULE1"). The compliance engine of the subject calls the method "VERIFY" on the resource "MyPatch" associated with the rule "RULE2"; assuming that the subject is not compliant to the "RULE2", a corresponding response is returned. The same operations are repeated for the resource "MyProgram" of the rule "RULE1".

The two pending rules "RULE2" and "RULE1" are then actually applied on the subject. For this purpose, the method "SET_STATE" is called on the resource "MyPatch"; since the resource "MyProgram" is not installed yet, the application of the rule "RULE2" fails. Conversely, when the method "SET_STATE" is called on the resource "MyProgram", the application of the rule "RULE1" succeeds in bringing the resource "MyProgram" to its target state ("Intalled").

The same operations are repeated for the still pending rule "RULE2". However, when the method "SET_STATE" is called again on the resource "MyPatch", the application of the rule "RULE2" now succeeds (since the pre-requisite "MyProgram=Installed" is now available).

The compliance engine of the subject calls the method "VERIFY" again on the resource "MyPatch", and receives a

response indicating that the subject is compliant to the rule "RULE2"; likewise, the compliance engine determines that the subject is compliant to the "RULE1" as well. The state variable indicating the compliance to all the rules ("RULE1" and "RULE2") is then transmitted to the membership repository.

Similar considerations apply to different scenarios, which relate to alternative applications or involve other rules and subjects. Moreover, in a different embodiment of the invention a description of minimal, specific dependencies for each rule is transmitted to the subject together with a corresponding package for the installation of the corresponding resource. Those dependencies are checked before the installation of the resource; in this way, any failing rules can be identified without the need to attempt their actual application.

The management environment described above implements an adaptive model, which can be compared to the behavior of the human society. In the human society a central authority (such as the parliament) enacts a series of laws for the citizens. Each law has a general character, and defines different provisions for corresponding categories of citizens. For example, a law establishes a series of tax rates, each one to be applied for a corresponding income range; moreover, a further law specifies that an income-tax return (specifying the details of the corresponding payment) must be filed with a competent authority.

The parliament publishes the laws using an official bulletin. Every citizen has the responsibility to remain up to date with the laws. In this way, the citizen can calculate the taxes to be paid applying the rate corresponding to his/her income, pay the resulting amount, and then file the income-tax return (filled in with the details of the payment). Should the citizen try filing the income-tax return before paying the taxes, that income-tax return would be refused by the competent authority. In this way, the parliament does not need to know any information about the citizens; at the same time,

the correct sequence of the operations to be performed is controlled by every citizen directly.

The proposed solution clearly distinguishes from the enforcement model implemented by the resource management environments known in the art. Referring again to the parallel with the civil society, a central authority (such as a gas company) collects information about the conditions of every user; the central authority then issues a series of measures, each one individually tailored to the condition of the corresponding user. For example, the gas company detects the consumption of every user through a meter reading; only after that operation has been completed, the gas company can issue a corresponding invoice (wherein the amount to be paid is calculated according to the consumption). In this way, the gas company must collect all the required information about the customers; in addition, for each customer the gas company must verify the completion of that operation before issuing the corresponding invoice.

In other words, the proposed model provides an autonomic computing implementation of the management environment. The term autonomic comes from an analogy to the autonomic central nervous system in the human body, which adjusts to many situations automatically without any external help. Likewise, in the proposed management environment each subject is able to configure, tune and repair itself, as well as anticipate and solve performance problems automatically; as the same time, any dependencies can be resolved automatically by the subjects themselves.

More generally, the present invention proposes an autonomic management method for a data processing system; the system includes a plurality of subject entities and one or more authority entities. The method is used for self-configuring the subject entities; each subject entity belongs to one or more of a plurality of categories. The authority entity publishes a plurality of rules; each rule defines a target state for a category. The target state of one

or more rules is a prerequisite for the target state of one or more further rules. Each subject entity retrieves the rules for the corresponding category. Each retrieved rule is then applied; the application of the retrieved rule fails when one or more corresponding prerequisite is not available on the subject entity or succeeds to configure the subject entity according to the target state specified in the retrieved rule otherwise. In the method of the invention, the application of each failed rule is repeated to configure the subject entity according to the target state specified in the failed rule when each corresponding prerequisite is available.

The solution of the invention provides a management method that implements a self-adaptive model. In this way, the subjects are no longer passive entities, but they actively participate in the management process.

The proposed method is able to resolve any dependencies among the management actions automatically.

Moreover, the devised solution removes the need to specify any information relating to a possible conditioning in the application of the rules.

As a result, the definition of any solution to be deployed in the system is strongly simplified.

It should be noted that this result can be achieved reusing components already available in known management environments (based on the enforcement model). Indeed, the method of the invention only changes the way these components are used, not their operation.

The preferred embodiment of the invention described above offers further advantages.

Particularly, a deadlock is detected when the number of pending rules does not decrease after each application.

In this way, any conflict in the definition of the rules can be readily identified (thereby preventing inconsistency conditions).

As a further enhancement, the number of pending rules is reset whenever further rules are retrieved on the subject.

The proposed feature completely de-couples the retrieving of the rules from their application.

A suggested choice for implementing the above-mentioned feature is to abort the process responsible of applying the
5 rules.

This method is very simple, but at the same time effective.

A way to further improve the solution is to verify that the subject is still compliant to all the rules at the end of
10 their application.

In this way, any infinite loop can be readily detected.

Advantageously, the solution according to the present invention is implemented with a computer program, which is provided as a corresponding product stored on a suitable medium.
15 Moreover, it should be noted that the component of the program running on each subject is suitable to be implemented separately and put on the market even as a stand-alone product.

Alternatively, the program is pre-loaded onto the hard-disks, is sent to the computers through a network
20 (typically the INTERNET), is broadcast, or more generally is provided in any other form directly loadable into the working memories of the computers. However, the method according to the present invention leads itself to be carried out even with a hardware structure (for example, integrated in a chip of
25 semiconductor material), or with a combination of software and hardware.

Naturally, in order to satisfy local and specific requirements, a person skilled in the art may apply to the solution described above many modifications and alterations
30 all of which, however, are included within the scope of protection of the invention as defined by the following claims.

CLAIMS

1. In a data processing system (100) including a plurality of subject entities (110) and at least one authority entity (105), an autonomic management method (300) for
5 self-configuring the subject entities, each subject entity belonging to at least one of a plurality of categories, wherein the method includes the steps of:

the at least one authority entity publishing (324) a plurality of rules each one defining a target state for a
10 category, the target state of at least one rule being a prerequisite for the target state of at least one further rule,

each subject entity retrieving (320-322,326) the rules for the corresponding at least one category,

15 applying (427-439) each retrieved rule to configure the subject entity according to the target state, the application of the retrieved rule failing when at least one corresponding prerequisite is not available on the subject entity, and

repeating (427-448) the application of each failed rule
20 to configure the subject entity according to the target state specified in the failed rule when all the corresponding prerequisites are available.

2. The method (300) according to claim 1, further including the step of:

25 detecting (445,451) a deadlock condition on the subject entity when the number of the corresponding failed rules does not decrease after each step of repeating the application.

3. The method (300) according to claim 2, further including the steps of:

30 the subject entity retrieving (320-322,326) further rules for the corresponding at least one category, and

resetting (415) the number of the corresponding failed rules in response to the retrieving of the further rules.

4. The method (300) according to claim 3, further including the step of:

aborting (330) the execution of a loop including the steps of applying the retrieved rules (427-439) and repeating the application of the failed rules (427-448) in response to the retrieving of the further rules.

5. The method (300) according to any claim from 1 to 4, further including the steps of:

the subject entity verifying (454-457) the compliance to each retrieved rule in response to the successful application of all the retrieved rules, and

detecting (460,466) a loop condition when the subject entity is not compliant to at least one retrieved rule.

6. A computer program (203;218,254), directly loadable into a working memory of a data processing system (100) including a plurality of subject entities (110) and at least one authority entity (105), for performing the method (300) of any claim from 1 to 5 when the program is run on the system.

7. In a data processing system (100) including a plurality of subject entities (110) and at least one authority entity (105), a computer program (218,254) directly loadable into a working memory of a computer of a subject entity for performing an autonomic management method (300) for self-configuring the subject entity when the program is run on the computer, each subject entity belonging to at least one of a plurality of categories and the at least one authority entity publishing (324) a plurality of rules each one defining a target state for a category, the target state of at least one rule being a prerequisite for the target state of at least one further rule, wherein the method includes the steps of:

retrieving (320,326) the rules for the corresponding at least one category,

applying (427-439) each retrieved rule rule to configure the subject entity according to the target state, the application of the retrieved rule failing when at least one corresponding prerequisite is not available on the subject
5 entity, and

repeating (427-448) the application of each failed rule to configure the subject entity according to the target state specified in the failed rule when all the corresponding prerequisites are available.

10 8. A program product (175) comprising a computer readable medium on which the program (203;218,254) of claim from 6 or 7 is stored.

9. In a data processing system including a plurality of subject entities (110) and at least one authority entity
15 (105), an autonomic management system (100) for self-configuring the subject entities, each subject entity belonging to at least one of a plurality of categories, wherein the at least one authority entity includes means (203) for publishing a plurality of rules each one defining a target
20 state for a category, the target state of at least one rule being a prerequisite for the target state of at least one further rule, and wherein each subject entity includes means (254) for retrieving the rules for the corresponding at least one category, means (254) for applying each retrieved rule to
25 configure the subject entity according to the target state, the application of the retrieved rule failing when at least one corresponding prerequisite is not available on the subject entity, and means (254) for repeating the application of each failed rule to configure the subject entity according to the
30 target state specified in the failed rule when all the corresponding prerequisites are available.

10. In a data processing system including a plurality of subject entities and at least one authority entity, a computer (110) of a subject entity suitable to be used in an autonomic management system (100) for self-configuring the subject
5 entity, each subject entity belonging to at least one of a plurality of categories and the at least one authority entity publishing a plurality of rules each one defining a target state for a category, the target state of at least one rule being a prerequisite for the target state of at least one
10 further rule, wherein the computer includes means (254) for retrieving the rules for the corresponding at least one category, means (254) for applying each retrieved rule to configure the subject entity according to the target state, the application of the retrieved rule failing when at least
15 one corresponding prerequisite is not available on the subject entity, and means (254) for repeating the application of each failed rule to configure the subject entity according to the target state specified in the failed rule when all the corresponding prerequisites are available.

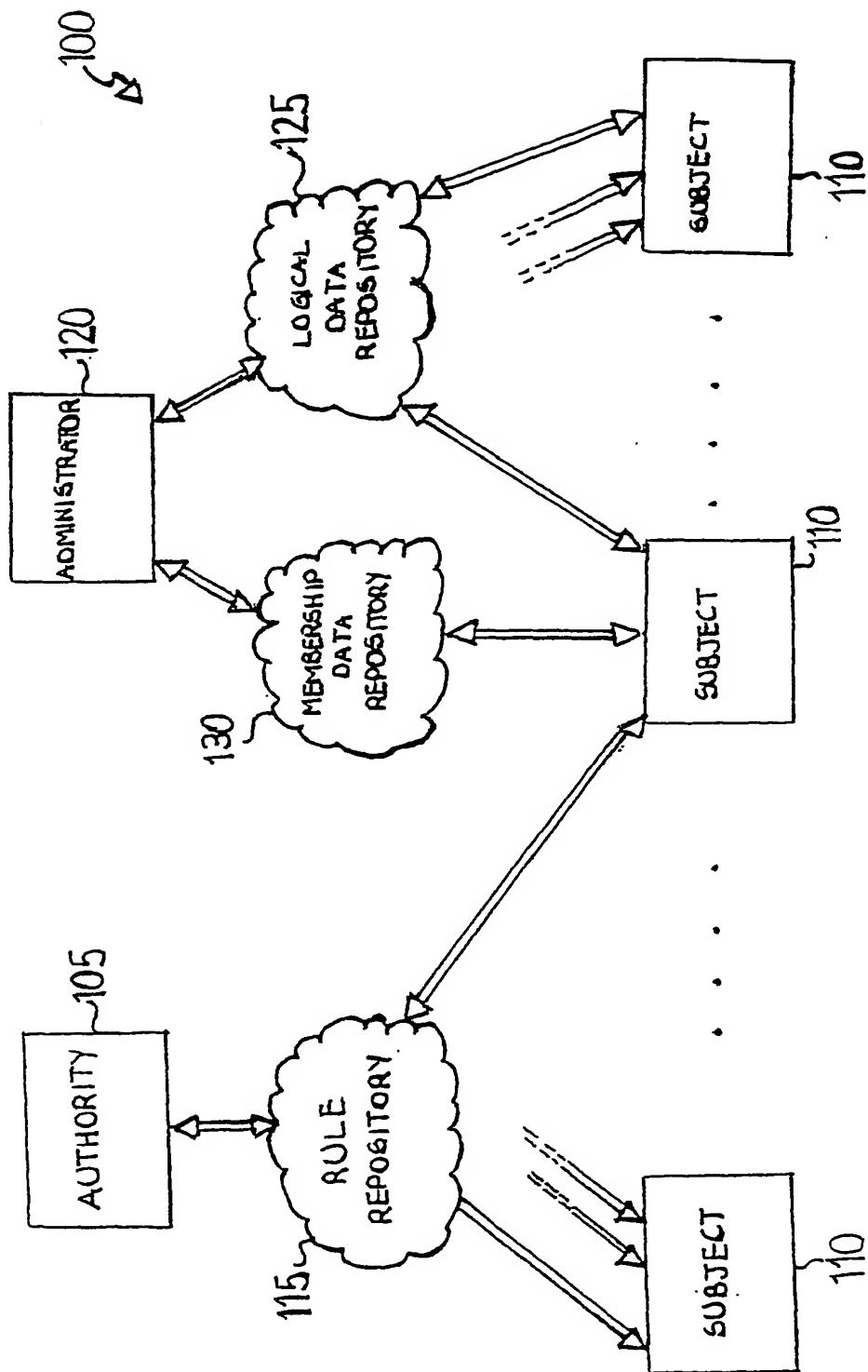


Fig. 1a

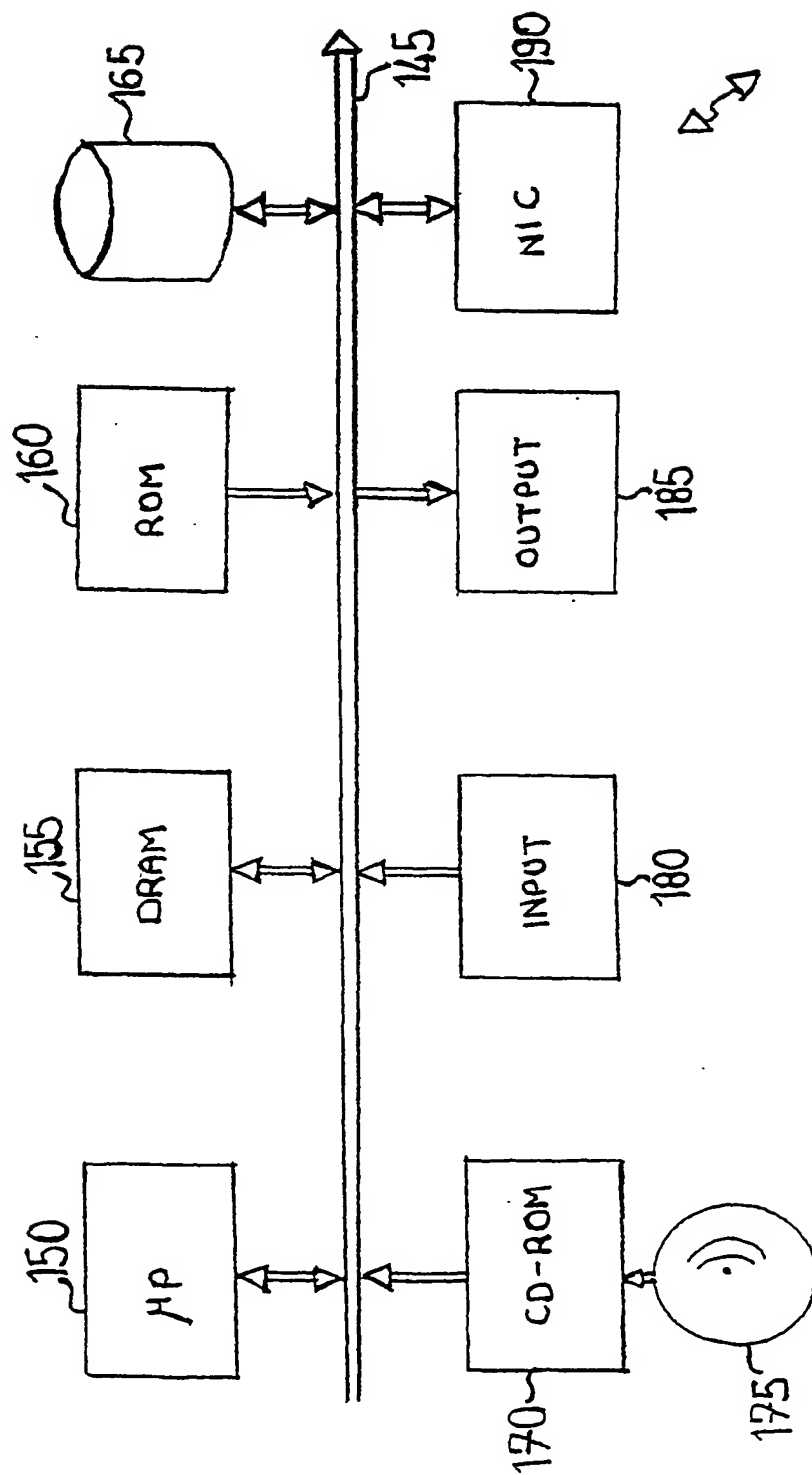


Fig. 1b

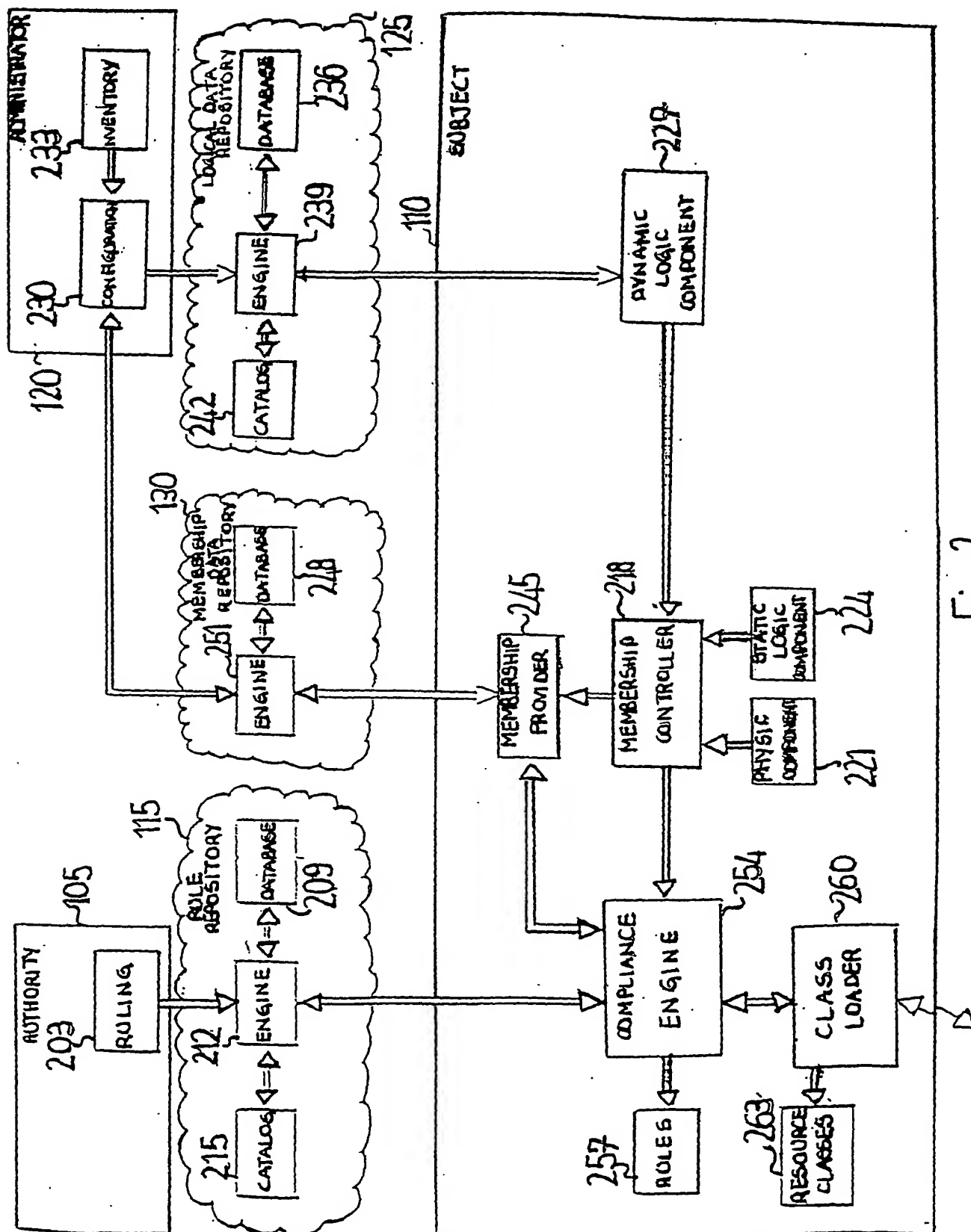
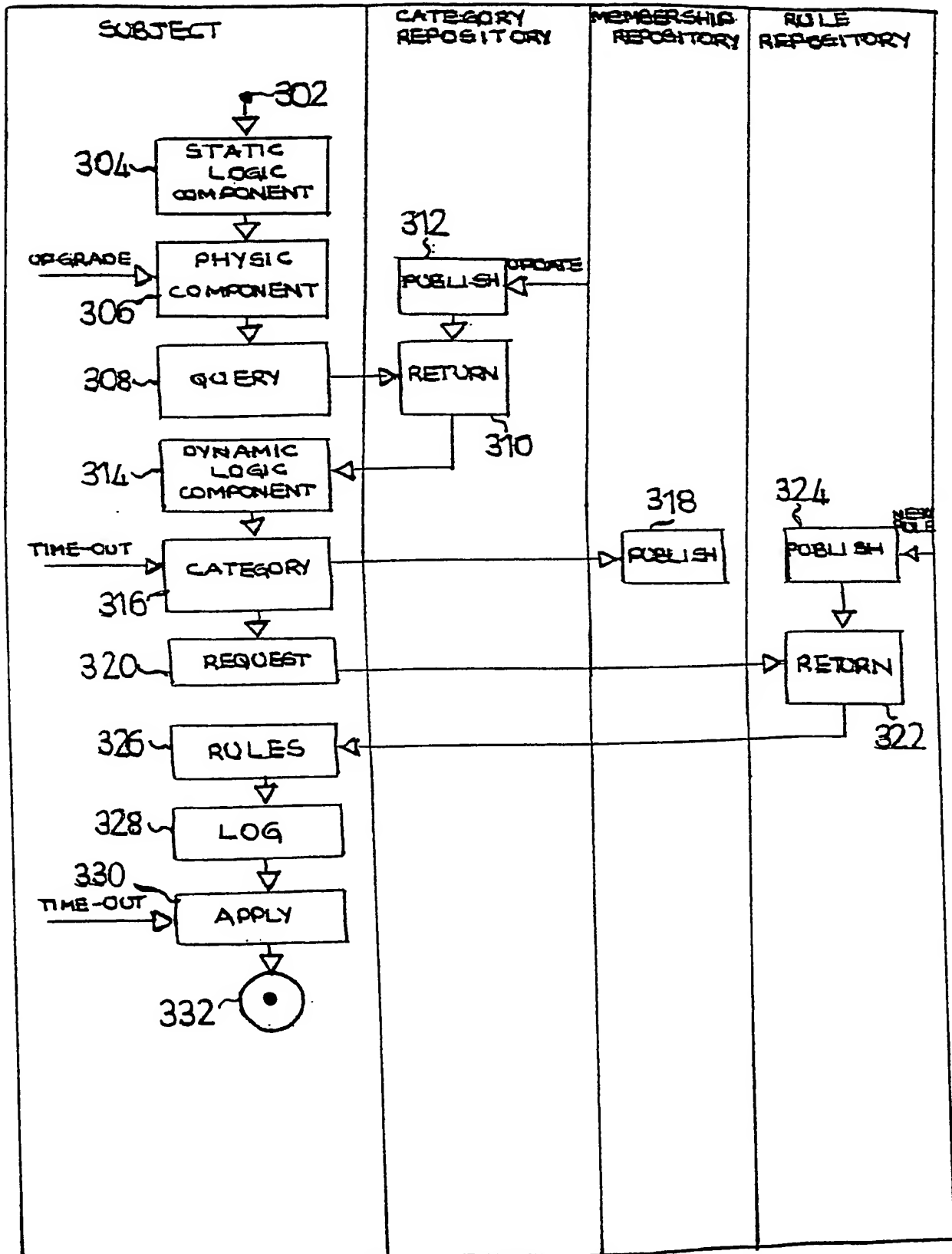


Fig.2



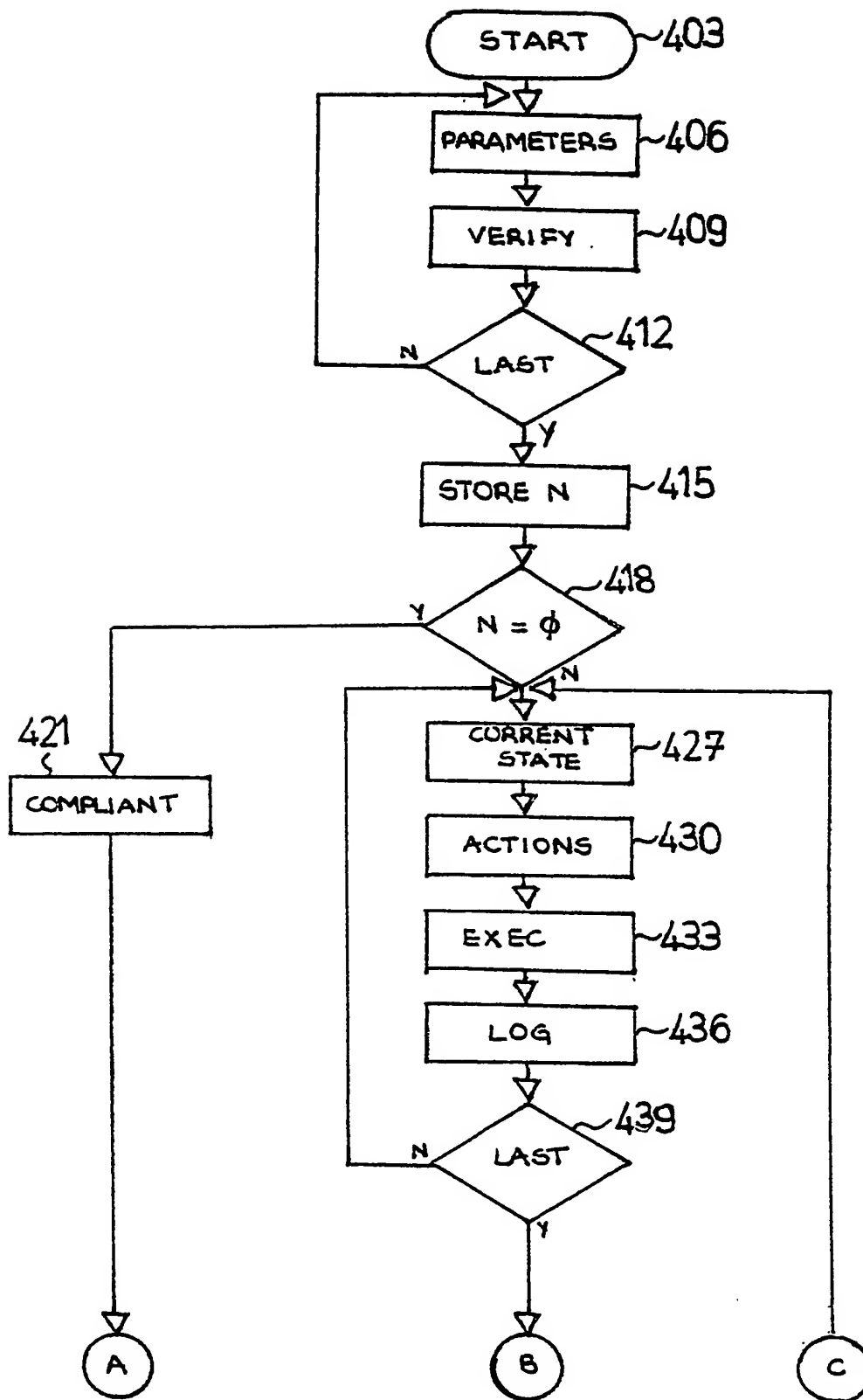


Fig. 4a

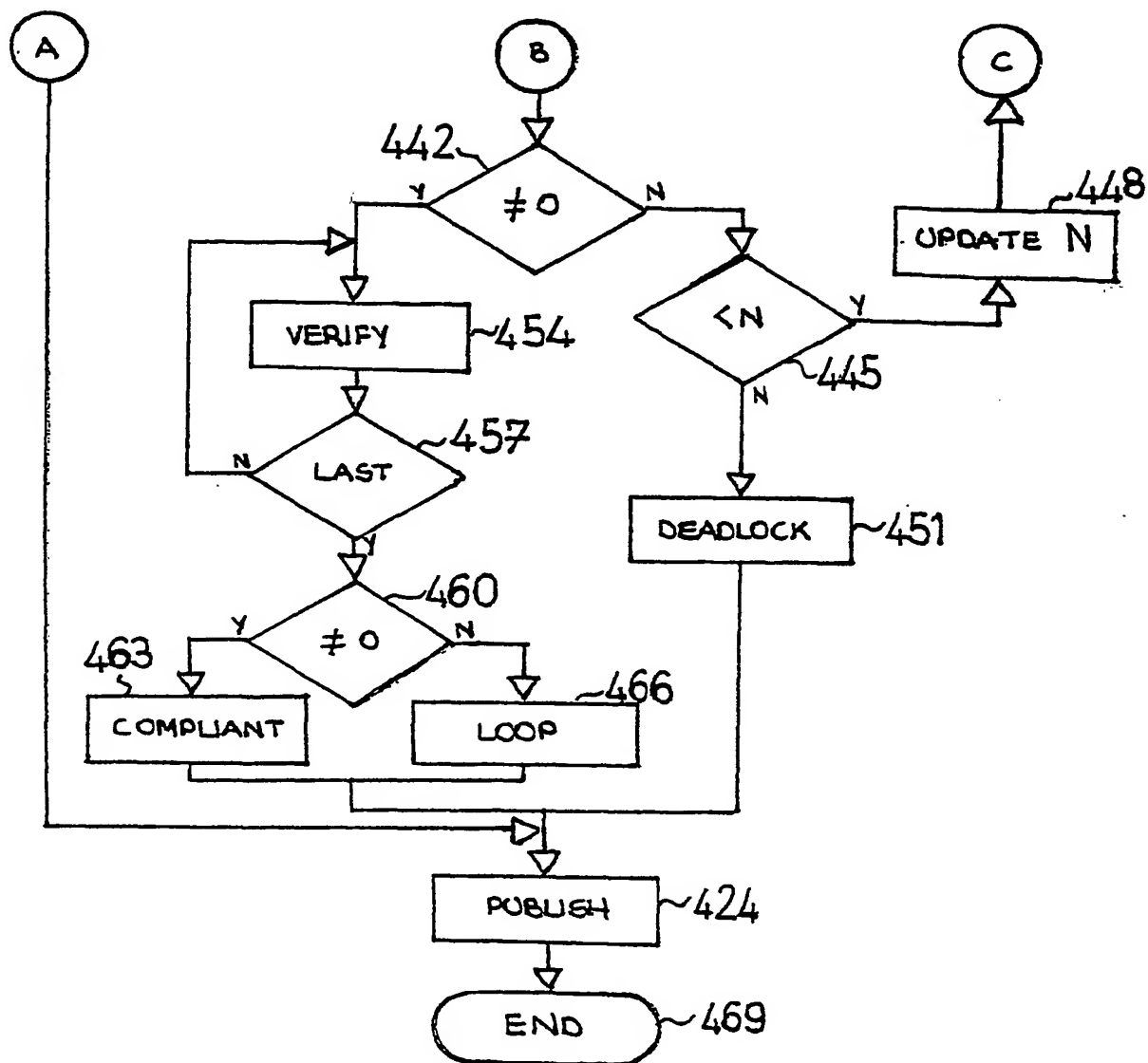


Fig. 4b

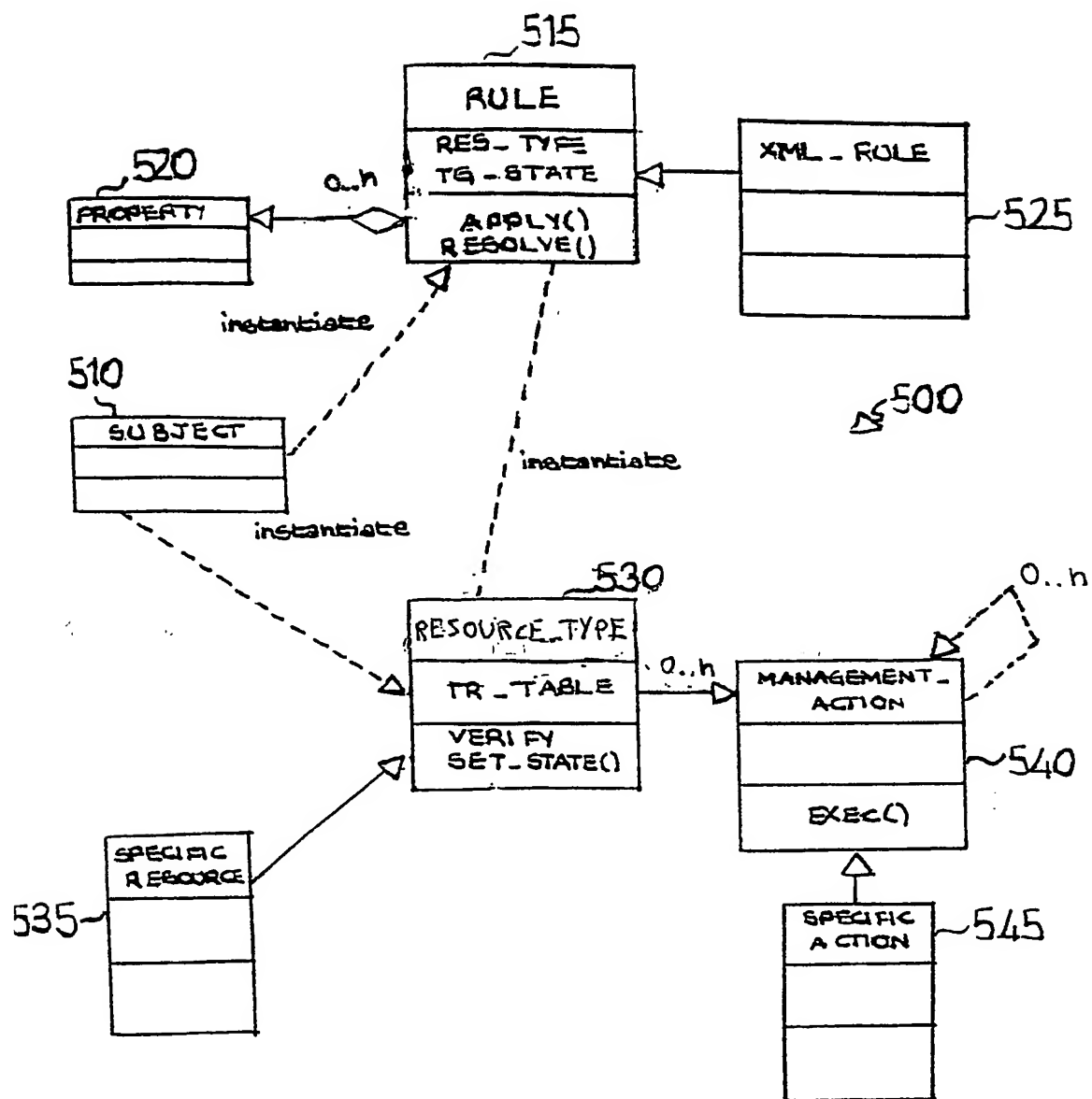


Fig. 5

